BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 199 301 A1

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 86105371.8

(22) Date of filing: 18.04.86

(51) Int. Cl.⁴: C07K 13/00 , C12N 15/00 , C12P 21/02 , C12P 21/00 , G01N 33/569 , A61K 39/21

(30) Priority: 19.04.85 US 725021

(43) Date of publication of application: 29.10.86 Bulletin 86/44

(84) Designated Contracting States: AT BE CH DE FR GB IT LI NL SE

(71) Applicant: F. HOFFMANN-LA ROCHE & CO. Aktiengesellschaft

CH-4002 Basel(CH)

Applicant: THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by the Secretary of the DEPARTMENT OF HEALTH AND HUMAN SERVICES

Washington, DC 20203(US)

(72) Inventor: Crowl, Robert Mitchell
115 Anderson Parkway
Cedar Grove, N.J. 07009(US)
Inventor: Gallo, Robert Charles
8513 Thornden Terrace
Bethesda Maryland 20817(US)
Inventor: Reddy, Eragam Premkumar
51 South Mountain Avenue
Montclair, N.J. 07042(US)
Inventor: Shaw, George Mead
2101 Inverness Cliffs
Birmingham Alabama 35243(US)
Inventor: Wong-Staal, Flossie Yeeching
18703 Ginger Court
Germantown Maryland 20874(US)

(74) Representative: Lederer, Franz, Dr. et al
Vanderwerth, Lederer & Riederer
Patentanwälte Lucile-Grahnstrasse 22
D-8000 München 80(DE)

EP 0 199 301 A1

(54) Recombinant acquired immune deficiency syndrome (AIDS) viral envelope protein and method of testing for AIDS.

(57) Recombinant envelope proteins of the etiologic agent of acquired immune deficiency syndrome - (AIDS) useful in diagnosis and therapy of AIDS and a method for their preparation are described. Proviral DNA is transferred into a host cell after engineering into an expression vector which produces the envelope protein. A method of testing human blood for the presence of antibodies to the AIDS virus using the AIDS envelope protein is also described.

lymphadenopathy-associated virus (LAV), AIDS-associated retrovirus (ARV) and the human T-cell leukemia/lymphoma/lymphotropic virus type III - (HTLV III).

A further aspect of this invention relates to a diagnostic method for testing human blood for the presence of antibodies to the env AIDS protein. This aspect of the invention overcomes the problems of all previously used blood tests for AIDS. One of the problems is the use of compositions to bind AIDS antibody which contain proteins or peptides which were not derived solely from the AIDS etiologic agent. A composition using homogeneous envelope AIDS protein of this invention overcomes the nonspecificity of the prior tests or assays. Yet another aspect of this invention is a diagnostic method for detecting and/or determining the presence of the antigen in human blood.

Another aspect of this invention is to use the env AIDS proteins of the instant invention as antigens suitable for providing protective immunity against AIDS when incorporated into a vaccine.

Brief Description of the Drawings

Fig. 1. The nucleotide sequence of the envelope gene of the HTLV-III proviral genome (HXB-3).

Fig. 2. Comparison of the amino acid sequence of the env protein of the five purported etiologic agents of AIDS. Amino acid sequences are aligned to give maximum homology.

Fig. 3. Construction of the pEV/env44-640 expression plasmids. The upper left panel shows a simplified restriction site map of the 3.15 Kb EcoRI-XhoI segment of the HTLV-III genome which contains the env coding region (cross-hatched arrow). The right panel shows the structure and pertinent sequences of the pEV-vrf plasmids. The solid black region represents the synthetic ribosome binding site sequences upstream of the ATG initiation codon (overlined). See Example 2 for a detailed description of the env expression plasmid constructions.

Fig. 4. Western blot analysis of env coded antigens produced in E. coli. Total bacterial proteins were resolved by SDS-PAGE, electro-blotted onto a nitrocellulose filter, and env encoded proteins were detected by reacting with human sera as described in Example 5: a) negative control, cells containing pJCL-E30 (p21T) induced at 42°C for 2 hours; b) uninduced control, cells containing pEV3/env44-640 maintained at 30°C; c) pEV3/env44-640; d) pEV1/env44-640; and e) pEV3/env205-640 induced at 42°C for 2 hours.

Fig. 5. Recognition of bacterially synthesized HTLV-III env gene products by antibodies in AIDS patient sera. Bacterial lysates containing recombinant env proteins were subjected to Western blot analysis as described in Example 5. Individual strips were then incubated with a 1000-fold dilution of individual sera followed by treatment with ^{125}I-labeled protein A. (upper part) Serum samples were from the following donors: (lane 1) normal healthy donor; (lanes 2-18) AIDS patient sera collected from the West Coast of the USA. - (Lower part) Serum samples were taken from the following donors: (lane 1) donor found to be HTLV-1(+) by Elisa using disrupted virus; (lanes 4, 5, 11 and 15) healthy, normal donors; (lanes 2, 3, 6, 8, 10, 12, 13, 14, 16, 17 and 18) AIDS patient sera from the East Coast of the USA.

Fig. 6A. The amino acid sequence of the AIDS envelope protein.

Fig. 6B. The amino acid distribution of the AIDS envelope protein.

Fig. 7. Construction of the expression vector pRC23. The Shine-Dalgarno sequence (SD) is overlined and the location of the synthetic ribosome binding site sequence in the plasmid is represented by the solid black segment. The plasmid contains the entire sequence of pBR322 and thus confers resistance to both ampicillin (amp^R) and tetracycline (tet^R).

Fig. 8. Construction of the pEV-vrf vectors. The synthetic oligonucleotides for each plasmid which were placed downstream of the SD sequence in pRC23 are shown with the locations of the restriction enzyme cleavage sites. The ATG initiation codon is overlined, and the placement of the additional A-T base pairs is designated by the rectangle. The plasmids confer resistance to ampicillin only.

cytopathic characteristic of HTLV III was critical to determining ultimately the specific retroviral origin of the disease. Thus the etiologic agent of AIDS was isolated by use of immortalized human neoplastic T cell lines (HT) infected with the cytopathic retrovirus characteristic of AIDS, isolated from AIDS afflicted patients. Seroepidemiological assays using this virus showed a complete correlation between AIDS and the presence of antibodies to HTLV III antigens [Sarngadharan, M.G. et al., supra; Schüpbach, J. et al., supra]. In addition, nearly 85% of patients with lymphadenopathy syndrome and a significant proportion of asymptomatic homosexual men in AIDS endemic areas were also found to carry circulating antibodies to HTLV III. Taken together, all these data indicate HTLV III to be the etiologic agent for AIDS.

Until the successful culturing of AIDS virus using H-9 cell line [PCT application, publication no. WO 85/04897] the env AIDS protein of the AIDS virus had not been isolated, characterized or synthesized. This in major part is due to the fact that the virus is cytopathic and thus isolation of the virus was not possible [Popovic, M. et al., supra]. Once the human T-cell line resistant to the cytopathic effects of the virus was discovered, a molecular clone of proviral DNA could be achieved.

The need for a sensitive and rapid method for the diagnosis of AIDS in human blood and its prevention by vaccination is very great. Virtually all the assays/tests presently available are fraught with errors. In fact the Center for Disease Control - (CDC) has indicated that presently available tests be used solely for screening units of blood for antibody to HTLV III. The CDC went further by stating that the presently available ELISA tests can not be used for general screening of high risk pupulations or as a diagnostic test for AIDS - [Federal Register 50(48), 9909, March 12, 1985]. The errors have been traced to the failure to use a specific antigenic protein of the etiologic agent for AIDS. The previously used proteins were derived from a viral lysate. Since the lysate is made from human cells infected with the virus, i.e. the cells used to grow the virus, the lysate will contain human proteins as well as viral proteins. Thus preparation of a pure antigen of viral protein is very difficult. The antigen used produced both false positive and false negative results [Budiansky, S., "AIDS Screening, False Test Results Raise Doubts", Nature 312, 583(1984)]. The errors caused by the use of such lysate proteins/peptides can be avoided by using a composi tion for binding AIDS antibodies which is substantially free of the non-AIDS specific proteins. Compositions that are substantially pure AIDS envelope protein can be used as antigens.

The AIDS envelope protein of the instant invention has been established to have conserved epitopes which permit its use to screen for, diagnose and/or prevent by vaccination the infection by AIDS virus. The instant invention demonstrates that the envelope protein with its conserved epitopes includes all the variants which have been claimed as the sole etiologic agent.

The envelope AIDS protein of the present invention may be produced by conventionally known methods. The processes by which the novel protein may be produced can be divided into three groups: (1) chemical synthesis; (2) preparation of a gene prepared by chemical synthesis which is inserted into a host and a protein is produced by the host; and (3) a corresponding gene obtained biotechnically is inserted into a host and a protein is produced by the host.

In one embodiment of this invention, recombinant DNA techniques are utilized by which env AIDS DNA from a natural source is introduced into a cell to produce the env AIDS protein. One method of obtaining DNA which encodes env AIDS is to read the genetic code in reverse and synthesize an oligodideoxynucleotide which should encode the env AIDS amino acid sequence. As the env protein has not been isolated or characterized this approach cannot be pursued.

Alternatively gene expression can be obtained using recombinant DNA technology if DNA isolated from natural sources is used instead of synthetic DNA.

Summary of the Invention

This invention is directed to the engineering of HTLV III env gene into suitable expression vectors; transformation of host organisms with such expression vectors; and production of envelope AIDS protein (env AIDS) by culture of such transformed cells. Another aspect of the present invention relates to the isolation and use of the resulting recombinant env AIDS protein.

Another aspect of the present invention is the identification and determination of the proviral DNA sequence. More specifically, this aspect of the invention relates to determination and comparison of the proviral nucleotide sequence of the envelope genes of the purported etiologic agent of AIDS i.e.

M = Methionine; N = Asparagine; P = Proline; Q = Glutamine; R = Arginine; S = Serine; T = Threonine; V = Valine; W = Tryptophan; Y = Tyrosine; Z = Glutamine or Glutamic Acid.

In accordance with the present invention, the search for the envelope protein of the etiologic agent for acquired immune deficiency syndrome (AIDS) has led to the isolation and sequencing of the proviral gene of the AIDS virus. It has now been discovered, for what is believed to be the first time that the postulated etiologic agents of AIDS, lymphadenopathy-associated virus (LAV), AIDS-associated retrovirus (ARV) and human T-cell leukemia/lymphoma/lymphotropic virus (HTLV III) are in fact variants of the same virus. For purposes of this invention, in the specification and claims the virus causing AIDS will be referred to herein as AIDS virus. AIDS virus will be understood to include the variants which have been postulated as the causative agents of AIDS, namely LAV, ARV and HTLV III. The envelope protein of the AIDS virus (env AIDS) is a 97,200 dalton protein with 32 potential N-glycosylation sites. Nucleotide sequence analysis of the AIDS envelope gene of the putative etiologic agents of AIDS demonstrates that all the viruses are variants of the same virus. That means that there is approximately 1 to 20% divergence or variation from the sequence of the envelope gene of HTLV III and the sequences of the envelope genes of the other viruses LAV and ARV-2. The amino acid sequence of the env AIDS is set forth in Figure 6(a). The amino acid distribution is set forth in Figure 6(b).

The nucleotide sequence of the envelope gene is shown in Figure 1. The proviral DNA sequence, using methods known to one of ordinary skill in the art such as the chemical degradation method of Maxam and Gilbert of the M13 sequencing system of Messing which is a modification of the dideoxy nucleotide chain termination method of Sanger, was analyzed to determine the location of the region coding for the envelope protein. The location of an open reading frame, i.e. a long stretch of triplet codons not interrupted by a translational stop codon, for the envelope gene was determined. The open reading frame coding for the env gene is 863 amino acids and contained an ATG codon at the eighth position from the 5' end of the reading frame. The ATG codon is known to be a universal translation-initiation codon.

The integrated proviral genome of HTLV-III was cloned from the genomic DNA of H9 cells infected with HTLV-III [Shaw, G.M. et al., "Molecular characterization of Human T-cell leukemia - (lymphotropic) virus type III in the acquired immune deficiency syndrome", Science 226, 1165-1171 (1984)]. Since the HTLV-III provirus was found to lack XbaI restriction sites, a genomic library was constructed by using XbaI digested H9/HTLV-III DNA. There are several methods available to one of ordinary skill in the art for screening the bacterial clones containing the AIDS env protein cDNA. These include, for example, RNA selection hybridization, differential hybridization with a synthetic probe or screening for clones that produce the desired protein by immunological or biological assays. From the genomic library, colonies of cells transformed with DNA that contains the HTLV III sequences were selected by hybridization screening of the library with HTLV III cDNA. The DNA insert of the hybridization-positive clone, HXB-3, was excised from the plasmid DNA and sequenced.

The predicted product of the env gene shares many features in common with the envelope gene products of other retroviruses. Thus, a hydrophobic region is seen in the middle of the protein (amino acids 519-534) which includes a processing site for the cleavage of the precursor protein into exterior and transmembrane proteins. Similarly, the amino terminal end contains a short stretch of hydrophobic amino acids (amino acids 17-37) which constitutes a potential signal sequence. The HTLV-III envelope precursor differs from the other retroviral envelope protein precursors in that it contains an additional stretch of 180 amino acids at the carboxy terminus.

Polymorphism within the Envelope Region of AIDS Virus

The recent publication of the nucleotide sequences of LAV, ARV-2 and HTLV-III [Ratner, L. et al., supra; Sanchez-Pescadon, R., et al., supra; Wain-Hobson, S., et al., supra] allows a detailed comparison of these various isolates obtained from AIDS patients from different parts of the world. HTLV-III clones were isolated from AIDS patient lymphocytes obtained from the east coast of the United States, while LAV was isolated from a French man and ARV was isolated from a patient in California. A comparison of the sequence data confirms the earlier observations made using restriction enzyme site analysis which showed approximately 10% variation. The present analysis shows that the various isolates show the greatest amount of conservation in the gag and pol regions while the most divergence occurs in the env region. A comparison of the five env sequences is presented in Figure 2. With respect to the envelope gene, HTLV-III and LAV are more closely related to each other than the ARV clone. Approximately 1.6%

Detailed Description of the Invention

In the description the following terms are employed:

Nucleotide: A monomeric unit of DNA consisting of a sugar moiety (pentose), a phosphate, and either a purine or pyrimidine base (nitrogenous heterocyclic). The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose). That combination of a base and a sugar is called a nucleoside. Each nucleotide is characterized by its base. The four DNA bases are adenine ("A"), guanine ("G"), cytosine ("C") and thymine ("T").

DNA Sequence: A linear array of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

Codon : A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translation start signal or a translation termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"). TAG, TAA and TGA are translation stop signals and ATG is a translation start signal.

Reading Frame: The grouping of codons during translation of mRNA into amino acid sequences. During translation the proper reading frame must be maintained. For example, the sequence GCTGGTTGTAAG may be translated in three reading frames or phases, each of which affords a different amino acid sequence:

GCT GGT TGT AAG = Ala-Gly-Cys-Lys

G CTG GTT GTA AG = Leu-Val-Val

GC TGG TTGTAA G = Trp-Leu-(STOP)

Polypeptide: A linear array of amino acids connected one to the other by peptide bonds between the α-amino and carboxy groups of adjacent amino acids.

Genome: The entire DNA of a cell or a virus. It includes inter alia the structural genes coding for the polypeptides of the substance, as well as operator, promoter and ribosome binding and interaction sequences, including sequences such as the Shine-Dalgarno sequences.

Structural Gene : A DNA sequence which encodes through its template or messenger RNA ("mRNA") a sequence of amino acids characteristic of a specific polypeptide.

Transcription: The process of producing mRNA from a structural gene.

Translation: The process of producing a polypeptide from mRNA.

Expression: The process undergone by a structural gene to produce a polypeptide. It is a combination of transcription and translation.

Plasmid : A circular double-stranded DNA molecule that is not a part of the main chromosome of an organism containing genes that convey resistance to specific antibiotics. When the plasmid is placed within a unicellular organism, the characteristics of that organism may be changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet^R) transforms a cell previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

Cloning Vehicle : A plasmid, phage DNA or other DNA sequences which are able to replicate in a host cell, which are characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which contain a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is often called a vector.

Cloning: The process of obtaining a population of organisms or DNA sequences derived from one such organism or sequence by asexual reproduction.

Recombinant DNA Molecule or Hybrid DNA: A molecule consisting of segments of DNA from different genomes which have been joined end-to-end outside of living cells and have the capacity to infect some host cell and be maintained therein.

The nomenclature used to define the peptides or proteins is that used in accordance with conventional representation such that the amino group at the N-terminus appears to the left and the carboxyl group at the C-terminus to the right. By natural amino acid is meant one of the amino acids commonly occurring in natural proteins comprising Gly, Ala, Val, Leu, Ile, Ser, Thr, Lys, Arg, Asp, Asn, Glu, Gln, Cys, Met, Phe, Tyr, Pro, Trp and His. By Nle is meant norleucine, and by Nva is meant norvaline. Where L and D forms are possible, it is the L-form of the amino acid that is represented unless otherwise expressly indicated. In addition, amino acids have been designated by specific letters of the alphabet such that: A = Alanine; B = Aspartic Acid or Asparagine; C = Cysteine; D = Aspartic Acid; E = Glutamic Acid; F = Phenylalanine; G = Glycine; H = Histidine; I = Isoleucine; K = Lysine; L = Leucine;

tion of an appropriate host is also controlled by a number of factors recognized by the art. These include, for example, compatibility with the chosen vector, toxicity of proteins encoded by the hybrid plasmid, ease of recovery of the desired protein, expression characteristics, biosafety and costs. A balance of these factors must be struck with the understanding that not all hosts may be equally effective for expression of a particular recombinant DNA molecule.

A preferred embodiment of the instant invention is to express segments of the AIDS env protein in E. coli by inserting restriction fragments isolated from the cloned proviral genome into the versatile pEV-vrf (variable reading frame) expression plasmids (for details of construction see Example 2). These versatile pEV-vrf plasmids are derivatives of pBR322 which contain the phage lambda P_L promoter, a synthetically-derived ribosome-binding site, and convenient cloning sites (EcoRI, BamHI, ClaI and HindIII) just downstream to the initiation codon (Figure 8). A set of three plasmids was constructed to accomodate all three translational reading frames. The P_L promotor is regulated by a temperature-sensitive cI repressor encoded on the compatible plasmid pRK248cIts [ATCC 33766; Bernard, H.U. and Helinski, D.R., "The use of the λ phage promotor P_L to promote gene expression in hybrid plasmid cloning vehicles", Meth. Enzymol. 68, 482-492 (1979)]. These expression plasmids have been used to produce substantial amounts of several heterologous proteins in E. coli including v-bas p21 [Lacal, J.C. et al., "Expression of Normal and Transforming H-ras genes in E. coli and purification of their encoded p21 proteins", PNAS 81, 5305-5309 (1984)] and murine interleukin-1 [Lomedico, P.T. et al., "Cloning and Expression of Murine Interleukin-1 cDNA in E. coli", Nature 312, 458-462 (1984)].

In the present synthesis the preferred initial cloning vehicle is the bacterial plasmid pBR322 - (ATCC 37017) and the preferred initial restriction endonuclease sites therein are the EcorRI and HindIII sites (Figure 3). Insertion of proviral DNA contained within the genome of H9 cells into these sites provides a large number of bacterial clones each of which contains one of the proviral DNA genes or fragments thereof present in the genome of H9 cells. Only a very few of these clones will contain the gene for env AIDS or fragments thereof.

The preferred host for initial cloning and expression of the env AIDS gene in accordance with this invention is E. coli MC 1061 [Casadaban, M.J. and Cohen, S.M., "Analysis of Gene Control Signals by DNA Fusion and Cloning in E. coli", J. Mol. Biol., 138, 179-207 (1980)].

The coding sequences for amino acid residues #44 to 640 of the env protein are located downstream of the P_L promoter between the KpnI and HindIII sites on the restriction map as shown in Figure 3. Aside from the location of these convenient restriction sites, these sequences were chosen for bacterial expression experiments because they did not include the amino-terminal signal peptide as well as the hydrophobic transmembrane segment at the carboxyl end. These sequences were excluded to avoid possible toxicity problems which can occur when hydrophobic proteins are over-produced in bacterial cells. In a preferred embodiment of this invention an expression plasmid was constructed that would direct the synthesis of this segment of the env gene product (designated pEV/env 44-640), an intermediate construction was first made by inserting a 2400 bp EcoRI-HindIII fragment between the EcoRI and HindIII sites in the pEV-vrf plasmids. The HTLV-III sequences (600 bp) between the EcoRI and the KpnI site were then removed from the intermediate construction as shown in Figure 3. These plasmid constructions were carried out with all three pEV-vrf plasmids so that subsequent deletions could be made and the correct reading frame maintained. In addition, the constructions made in the incorrect reading frames served as important controls in the expression experiments described below.

In another embodiment of this invention, a second set of expression plasmids were constructed in a similar fashion by deleting sequences between EcoRI and StuI sites which occur 483 bp downstream of the env gene. Again these deletions - (designated pEV/env 205-640) were made in all three reading frames. The translation termination codon used in all of the env expression plasmids is presumably an in-frame TAA located 23 bp downstream of the HindIII site in the plasmid. Thus, 8 amino acid residues at the carboxyl terminus are encoded by pBR322 sequences contained within the pEV-vrf expression plasmids.

Expression of ENV AIDS

There are several approaches to screen for bacterial clones containing env AIDS cDNA. These include, for example, RNA selection hybridization, differential hybridization, hybridization with a synthetic probe and screening for clones that produce

divergence was observed between the HTLV-III (HXB-3) and LAV sequence. Among the HTLV sequences, the divergence was about 1.6%. However, approximately 17% divergence was observed between HTLV-III and ARV-2 and this was more pronounced in the extracellular region of the envelope gene product (Figure 2). This high rate of divergence could be due to the geographical location from where the two isolates were derived or the time of isolation of these variants. ARV-2 was isolated from the west coast of the United States more recently. The HTLV-III isolates for which the nucleotide sequences have been determined were all obtained from the east coast of the United States a year earlier. LAV was obtained from a French patient who appears to have acquired the virus in New York about the same period. The observed differences in the sequence probably reflect divergent evolution of strains separated in time or geography or both. Within the env region, the highest level of divergence is in the extracellular portion of the protein.

Expression Vector

A wide variety of host/cloning vehicle combinations may be employed in cloning the double-stranded DNA. For example, useful cloning vehicles may consist of segments of chromosomal, nonchromosomal and synthetic DNA sequences, such as various known bacterial plasmids, e.g. plasmids from E. coli such as pBR322, phage DNA, and vectors derived from combinations of plasmids and phage DNAs such as plasmids which have been modified to employ phage DNA or other expression control sequences or yeast plasmids. Useful hosts may include microorganisms, mammalian cells, plant cells and the like. Among them microorganisms and mammalian cells are preferably employed. As preferable microorganisms, there may be mentioned yeast and bacteria such as Escherichia coli, Bacillus subtilis, Bacillus stearothermophilus and Actinomyces. The above-mentioned vectors and hosts may also be employed for the production of a protein from a gene obtained biologically as in the instant invention. Of course, not all host/vector combinations may be equally efficient. The particular selection of host/cloning vehicle combination may be made by those of skill in the art after due consideration of the principles set forth without departing from the scope of this invention.

Furthermore, within each specific cloning vehicle, various sites may be selected for insertion of the double-stranded DNA. These sites are usually designated by the restriction endonuclease which cuts them. For example, in pBR322 the EcoRI site is located just outside the gene coding for ampicillin resistance. Various sites have been employed by others in their recombinant synthetic schemes. Several sites are well recognized by those of skill in the art. It is, of course, to be understood that a cloning vehicle useful in this invention need not have a restriction endonuclease site for insertion of the chosen DNA fragment. Instead, the vehicle could be joined to the fragment by alternative means.

The vector or cloning vehicle and in particular the site chosen therein for attachment of a selected DNA fragment to form a recombinant DNA molecule is determined by a variety of factors, e.g., number of sites susceptible to a particular restriction enzyme, size of the protein to be expressed, susceptibility of the desired protein to proteolytic degradation by host cell enzymes, contamination of the protein to be expressed by host cell proteins difficult to remove during purification, expression characteristics, such as the location of start and stop codons relative to the vector sequences, and other factors recognized by those of skill in the art. The choice of a vector and an insertion site for a particular gene is determined by a balance of these factors, not all selections being equally effective for a given case.

There are several known methods of inserting DNA sequences into cloning vehicles to form recombinant DNA molecules which are equally useful in this invention. These include, for example, direct ligation, synthetic linkers, exonuclease and polymerase-linked repair reactions followed by ligation, or extension of the DNA strand with DNA polymerase and an appropriate single stranded template followed by ligation.

It should, of course, be understood that the nucleotide sequences of the DNA fragment inserted at the selected site of the cloning vehicle may include nucleotides which are not part of the actual structural gene for the desired polypeptide/protein or may include only a fragment of the complete structural gene for the desired protein. It is only required that whatever DNA sequence is inserted, a transformed host will produce a protein/peptide having an immunological activity to the AIDS env protein or that the DNA sequence itself is of use as a hybridization probe to select clones which contain DNA sequences useful in the production of polypeptides/proteins having an immunological activity to the AIDS env protein.

The cloning vehicle or vector containing the foreign gene is employed to transform a host so as to permit that host to express the protein or portion thereof for which the hybrid DNA codes. The selec-

tion. No reaction was observed with sera from healthy individuals or from HTLV-I infected individuals. The patient sera were derived from all parts of the United States including California and all AIDS patients' sera tested so far were found to be positive. The results suggest that these antibodies are mainly directed against the protein back-bone of the molecule.

It appears, therefore, that the env gene products constitute the best diagnostic reagents for the detection of AIDS associated antibodies. The env gene product of the instant invention encompasses a large portion of the protein molecule and contains both the conserved and divergent portions of the molecule. In spite of the divergence observed between HTLVIII and ARV-2 sequences the recombinant env proteins of the instant invention synthesized by the bacteria react with AIDS patient sera derived from both geographical locations of the United States. One hundred percent (100%) of AIDS patient sera (50 individual samples, 25 derived from the East Coast of the United States and 25 derived from California) tested showed high reactivity. This is strong evidence for the presence of conserved epitopes within the molecule against which the immune system could mount an antibody reaction. The human immune system may thus be mounting an immune response against conserved epitopes of the envelope molecule, as suggested by the reactivity of the AIDS patient sera. The observed divergence between various isolates of HTLV-III thus may not pose a problem for the use of recombinant protein as a vaccine. The 68Kd protein is ideally suited for such a purpose since it encompasses a large portion of the gene product and has the unique structural feature of containing both the extracellular hydrophilic region and the membrane associated hydrophobic regions. This structural feature makes it well suited for encapsulation into liposomes which have been used as vehicles for vaccination against other vital envelope proteins.

Based on these discoveries it is proposed that in the practice of screening blood for AIDS only AIDS envelope protein or a variant of said protein be utilized. Utilizing the env AIDS protein of the instant invention, human blood can be screened for the presence of antibodies to the AIDS virus. This and other techniques are readily determined, once, as taught for the first time by the present invention, the envelope AIDS protein has been recognized to be the envelope protein of the etiologic agent of AIDS. The foregoing and other objects, features and advantages of the invention will be apparent from the following examples of preferred embodiments of the invention.

Example 1

Molecular cloning and nucleotide sequence analysis of the HTLV-III proviral genome.

The integrated proviral genome of HTLV-III was recently cloned from the genomic DNA of H9 cells infected with HTLV-III [Shaw, G.M. et al., supra]. The proviral genome which was obtained by using XbaI digested H9/HTLV-III DNA contained two internal EcoRI sites within the viral genome and two additional sites in the cloning vector λ J1. These sites were used for further subcloning of the three DNA fragments of 5.5Kb, 4.5Kb and 1.1Kb into pBR322 (ATCC No. 37017). Nucleotide sequence analysis of the proviral genome was determined by the chemical degradation method of Maxam, A.M. and Gilbert, W., "Sequencing end-labelled DNA with base-specific chemical cleavages", Meth. Enzymol. 65, 499-560 (1980). For the sequence analysis, DNA inserts from the three subclones were isolated by electroelution and further cleaved with appropriate restriction enzymes. The DNA fragments were labelled at their 5'ends with γ-32P-ATP using polynucleotide kinase, or at their 3' ends with α-32P-NTP by filling in with DNA polymerase I - (Klenow fragment). The DNA fragments labelled at the two ends were cleaved with a second enzyme and the fragments labelled at a single end were purified on 5% acrylamide gels and used for sequence analysis. For the sequence analysis of the env gene, a shotgun approach was utilized where the 4.5 EcoRI fragment was cleaved with one of the following enzymes: BglII, HindIII, XhoI, AvaII, HinfI and Sau3A and the restriction fragments labelled and sequenced as described above. The nucleotide sequence of the envelope gene used in the present invention is shown in Figure 1.

Example 2

Construction of pEV/env 44-640

pRC2 is a derivative of pBR322 containing a unique Bgl II site adjacent (on the amp^R side) to the EcoRI site in the plasmid. This plasmid was constructed in the following manner. 20 µg of pBR322 plasmid DNA were digested with EcoRI and then split into two reactions. In one, the protruding 5' single-stranded termini were removed with S1 nuclease; in the other reaction, the termini were filled-in by inorporating deoxynucleotides with the Klenow fragment of DNA polymerase I. Both reactions were terminated by phenol extraction fol-

the desired protein by immunological or biological assays. Two methods are available to screen using immunological assay: screening of bacterial colonies for the presence of protein using antibody; and, preferably, the bacterial lysates are electrophoresed, blotted onto a nitrocellulose paper and then probed with the antibody.

In a preferred embodiment of this invention, cultures of the E. coli strain MC 1061 transformed with pRK248clts and the pEV 1, 2, or 3/env 44-640 (or pEV 1, 2 or 3/env 205-640) were grown in M9 medium at 30°C to mid-log phase and then induced by shifting to 42°C for 2 hr. Samples of the bacterial cultures were then taken and subjected to SDS-polyacrylamide gel electrophoresis, followed by Western blot analysis to detect env proteins. The protein blots were treated with antisera to env AIDS proteins isolated either from immunized rabbits or from AIDS patients previously shown to contain high titer antibodies to AIDS antigens. This was followed by incubation with ^{125}I-labelled Staphylococcous aureus protein A, washing and autoradiography. Similar results were obtained with both sera except that the human serum was found to contain much higher titers of anti-HTLV-III antibodies and was devoid of all background reactivity with the E. coli proteins. For this reason human antibodies were used in all subsequent characterization.

Figure 4 shows the pattern of reactivity of the env AIDS proteins synthesized in bacteria (recombinant proteins) with anti-HTLV-III antibodies. The open reading frame in pEV3/env 44-640 encodes a protein that should migrate as a 68 Kd band on the gel. In fact, a 68 Kd band is observed in the lane corresponding to the induced cells containing pEV3/env 44-640 (lane C). However, in addition to the 68 Kd band, these cells synthesized proteins of 35 Kd, 25 Kd and 17 Kd which specifically cross-reacted with anti-HTLV-III antibodies. No HTLV-III cross-reacting bands are evident in the uninduced control (Lane b) or in a second negative control sample (Lane a) of induced cells containing a plasmid that directs the synthesis of v-bas p21 oncogene product (Lacal, J.C. et al., supra). The appearance of multiple bands synthesized from the env gene sequences was an unexpected result. Another unexpected result was the synthesis of env gene products from the plasmid (pEV1/env 44-640) where the insert was placed in the wrong reading frame with respect to the initiator codon immediately down stream of the P_L promoter (Lane d). In this case, E. coli cells containing plasmid pEV1/env. 44-640 synthesized a 63 Kd protein in addition to the 35 Kd, 25 Kd and 17 Kd proteins. These results could be readily explained when the nucleotide sequence of the envelope gene (Fig. 1) was examined. About 155 bases downstream to the KpnI site is an ATG codon which appeared to be utilized for the synthesis of the env gene product by the expression plasmid pEV1/env 44-640. Internal translation initiation is also the likely explanation for the appearance of the 35Kd, 25Kd and 17Kd proteins. Initiation codons which are preceded by so-called Shine-Dalgarno sequences (AGGA) are found within the env coding region at locations that are consistent with the sites of the observed protein products.

To confirm the above interpretation and to rule out the possibility that the smaller proteins are not formed as a result of premature termination or from proteolytic cleavage of the larger product, another deletion mutant in which sequences between the KpnI and StuI sites were deleted were constructed. This expression plasmid contains the coding sequences from amino acid positions 205-640 which could code for a protein of 49 Kd. Analysis of the proteins induced from E. coli harboring this plasmid verified that, in fact, these cells synthesize a 49 Kd protein in addition to the 35 Kd, 25 Kd and 17 Kd proteins (lane e, Fig. 4). From these results, it was concluded that pEV3/env 44-640 expression plasmid directs the synthesis of a 68 Kd protein in addition to several additional smaller polypeptides (i.e., 35Kd, 25Kd and 17Kd) produced from all of the env expression plasmids resulting from internal translation initiation within the env gene.

Screening of AIDS SERA

Because anti-HTLV-III antibodies are found in more than 90% of the AIDS patients, it was of interest to see if the bacterially synthesized env gene products could be used as diagnostic tools for the detection of these antibodies. For this analysis, total cell protein from an induced bacterial culture was fractionated by SDS-PAGE and transferred to a nitrocellulose filter by Western blotting technique. Strips of the filter containing transferred proteins were reacted with 1000-fold diluted human sera, and the antigen-antibody complexes formed were detected by incubation of the strips with 125-I-labelled Staphylococus aureus protein A followed by autoradiography. Prominent bands corresponding to reaction of the antibody to the 68 Kd, 35 Kd, 25 Kd and 17 Kd proteins were consistently observed when the serum used was from patients with AIDS syndrome. The results of such assays with different human sera are presented in Figure 5. The negative controls used were normal human sera and serum from a patient with HTLV-I infec-

gested with StuI. Again, upon recircularization and blunt-end ligation, the EcoRI site was regenerated; however, an additional 483 bp of env coding sequences were removed.

Example 3

Bacterial Growth and Induction of env Gene Expression

Cultures of E. coli strain MC 1061 transformed with plasmid pRK248clts and the pEV1, -2, or -3/env plasmids were grown in M9 medium containing 0.5% glucose and 0.5% casamino acids at 30°C to mid-log phase and then induced by shifting to 42°C for 2 hr. The cells were collected by centri fugation and processed as described in Examples 4 and 5.

Example 4

Expression and Purification of Env AIDS

A homogeneous recombinant viral env AIDS was purified according to the following procedure. The env AIDS protein expressed by a microbe tends to associate with the membrane fractions of the host microbe, principally the inner membrane of the microbe. The following purification method was designed to deal with this finding.

This purification method comprises:

(a) lysis of transformed microbial cells producing recombinant env AIDS protein;

(b) separation of env AIDS associated cellular membranes from other cellular components;

(c) extraction of env AIDS from associated membranes; and

(d) chromatographic purification of the resultant extraction solution containing env AIDS to yield a substantially pure recombinant viral env protein.

More specifically, the preferred purification method for the preparation of substantially pure recombinant viral env protein comprises:

(a) cultivating a transformed organism containing a DNA sequence which codes for viral env protein;

(b) causing a culture of the transformed organism of step (a) to accumulate the env protein;

(c) lysing the culture of transformed organisms of step (b) to form a cell lysate mixture;

(d) isolating the cell membrane components of the cell lysate mixture of step (c);

(e) washing the isolated cell membrane components with an extraction solution to yield a wash solution containing env protein; and

(f) chromatographically purifying the wash solution of step (e) to yield a substantially pure env AIDS protein.

In carrying out this method it is preferred that the cells be lysed by sonification, although it is forseeable that other known methods such as enzyme or mechanical lysis could also be used. It is preferred that the cell membrane component, specifically the inner and outer membranes, be isolated from other cellular components by methods such as centrifugation. It has been found that env AIDS expressed by the transformed microorganism tends to become associated with the cellular membranes. Therefore, isolation of these membranes during the purification process ensures high purification levels and high purity env AIDS at the end of the purification procedure.

Once the cell membranes are isolated from the lysate mixture, they are washed with an extraction solution, preferably salt solutions and a detergent to yield a second solution containing approximately 50% env AIDS protein. Preferably the cell membranes are washed in four separate steps with the salt solutions and detergent although it is forseeable that certain of these steps could be combined, rearranged or eliminated. The first step of washing the cell membrane may be done with a salt solution, preferably 1M NaCl. In the second step the cell membrane is washed with a detergent solution, preferably 1% Triton X-100. In the third step, the cell membrane is washed with another salt solution, 1.75M to 3.5M guanidine HCl. The final wash is also with a salt solution preferably about 7M Guanidine HCl. The wash solution which results from the fourth and final wash comprises about 50% env AIDS.

The final 50% env AIDS wash solution is then further purified by a chromatography step, preferably reverse phase high performance liquid chromatography (HPLC). The HPLC step yields env AIDS protein in a substantially 100% pure

lowed by ethanol precipitation. Approximately 1 µg of DNA from each reaction was mixed with 90 pmoles of phosphorylated BglII linkers - (CAGATCTG, purchased from Collaborative Research) and incubated with T4 DNA ligase at 15°C for 18 hours. The ligation products were then digested with BglII and PstI and subjected to gel electrophoresis in 1% agarose. The 3600 bp and 760 bp fragments from both reactions were recovered from the gel. For the construction of pRC2, the 3600 bp from the Klenow reaction was ligated to the 760 bp fragment from the S1 reaction. To construct a plasmid with the BglII site on the other side of EcoRI (tet^R side), designated pRC1, the 3600 bp fragment from the S1 reaction was ligated to the 760 bp fragment from the Klenow reaction. E. coli strain RR1 (ATCC No. 31343) was transformed with the ligation mixtures, and transformants were selected on LB agar plates containing 50 µg/ml ampicillin. Transformants containing the expected plasmid constructions were identified by restriction analysis of the isolated plasmid DNA. DNA sequence analysis confirmed that the S1 nuclease treatment precisely removed the 5' single-stranded termini.

pRC23 (see Figure 7) was constructed by inserting into pRC2 a 250 bp BglII-HaeIII fragment containing the λ $P_L$promoter joined to a pair of complementary synthetic oligonucleotides comprising a model ribosome-binding site (RBS). The HaeIII site is located within the 5' non-coding region of the λN gene 115 bp downstream of the $P_L$transcriptional initiation site. Approximately 1 µg of a 450 bp BglII-HpaI fragment isolated from phage λ DNA was digested with HaeIII. 200 ng of the resulting digestion products were mixed with 60 pmoles each of phosphorylated synthetic oligonucleotides containing the model RBS. The ligated molecules were digested with BglII and EcoRI and separated on a 5% polyacrylamide gel. The 270 bp ligation product was recovered from the gel, mixed with gel purified pRC2 vector that had been digested with BglII and EcoRI, and incubated with T4 DNA ligase at 15°C for 15 hours. The ligation mixture was used to transform strain RR1(pRK248Clts). Transformants selected on ampicillin-containing medium were screened by restriction analysis of the isolated plasmid DNA. The expected plasmid construction, pRC23, was confirmed by further restriction enzyme digestions and by DNA sequence analysis across the EcoRI junction (Fig. 7).

For the construction of the pEV-vrf set of plasmids (see Figure 8), plasmid pRC23 was digested with EcoRI and HindIII and the pRC23/EcoRI-HindIII vector isolated by preparative agarose gel electrophoresis. The mixture of synthetic oligonucleotides (32, 33, and 34 nucleotides) was combined with the mixture of the complementary sequences, heated to 58°C for 5 minutes in 150 mM NaCl, and cooled slowly to allow annealing. 0.1 pmoles of the synthetic duplexes were added to 0.07 pmoles of the pRC23/EcoRI-HindIII vector and incubated with T4 DNA ligase at 15°C for 15 hours. Strain RR1 (λ cI857) was transformed with the ligation products. Six ampicillin resistant transformants were selected for DNA sequence analysis. Of the six, two contained the expected sequence for pEV-vrf1, one for pEV-vrf2, and three for pEV-vrf3 (Fig. 3).

For the expression of the AIDS env gene, one µg of a 2400 bp EcoRI -HindIII DNA fragment, which was isolated from the cloned HTLV-III proviral genome by preparative agarose gel electrophoresis, was mixed with 0.1 µg of EcoRI - HindIII digested vector DNA (pEV-vrf1, -2, or -3). After heating at 65°C for 3 minutes, the mixtures were chilled on ice, and 20 µl ligation reactions were assembled, containing 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 10 mM DTT, 0.3 mM ATP, and 200 units of T_4 DNA ligase. After incubation at 15°C for 4 hours, the reactions were terminated by heating at 65°C for 5 minutes. The ligation products were used to transform E. coli strain MC1061 containing plasmid pRK248clts. Transformants were selected on Luria broth agar containing 50 µg/ml ampicillin at 30°C for 18 hours. Plasmid DNA was isolated from 1 ml of each culture and subjected to restriction analysis. All 12 isolates contained the expected plasmid construction. These intermediate constructions were then used to make pEV1, -2, and -3/env 44-640 by deleting the 600 bp between the EcoRI and KpnI sites as described below.

Approximately 0.5 µg of plasmid DNA was digested with KpnI and EcoRI. The resulting termini were then treated with the Klenow fragment of DNA polymerase I in the presence of all four deoxyribonucleotides (at 100 µM) at 37°C for 30 minutes. This step results in the "filling-in" of the 5' overhang of the EcoRI terminus and the removal of the 3'overhang of the KpnI terminus. Upon recirculization of the linear plasmid and blunt-end ligation of these termini, an EcoRI site is regenerated. Transformants containing plasmids with the expected deletion were identified by restriction analysis.

A second set of deletion derivatives, designated pEV/env 205-640 was constructed in a similar fashion. A portion of the linear plasmid that had been digested with EcoRI and KpnI and treated with Klenow, as described above, was further di-

The size and shape of epitopes found in carbohydrate antigens have been extensively studied, but less is known about the structure of epitopes from protein molecules. Some epitopes of protein antigens have been defined at the level of their tertiary structure. In every instance, the epitopes were formed not by the primary sequences alone, but by the juxtaposition of residues brought together by the folding of the polypeptide chain(s) of the native molecule. In addition, the structure of the 68Kd env protein of the instant invention makes it particularly well suited for use as a vaccine. The 68Kd env protein comprises a large portion of the gene product which (a) was shown to be reactive with all the AIDS sera tested; and (b) has the unique structural feature of containing both an extracellular hydrophilic region and the transmembrane hydrophobic regions. The latter structural feature makes it well suited for use as a vaccine using liposome encapsulation to create a vehicle for administration.

Routes of administration, antigen dose, number and frequency of injections are all matters of optimization within the scope of ordinary skill in the art, particularly in view of the fact that there is experience in the art in providing protective immunity by the injection of other related antigens to provide immunity in other viral infections. It is anticipated that the principal value of providing immunity to AIDS infection will be for those individuals who have had no previous exposure to AIDS, e.g., individuals who are in the high risk population, such as homosexuals, drug addicts and people from Haiti and Central America and individuals who may be receiving blood transfusions. It is also anticipated that temporary immunity for infants may be provided by immunization of mothers during pregnancy.

Example 7

Diagnostic Test for AIDS

It is clear that the env gene proteins of the instant invention may be used as diagnostic reagents for the detection of AIDS-associated antibodies. It is also apparent to one of ordinary skill that a diagnostic assay for AIDS using polyclonal or monoclonal antibodies to the AIDS env protein of the instant invention may be used to detect the presence of the AIDS virus in human blood. In one embodiment a competition immunoassay is used where the antigenic substance, in this case the AIDS virus, in a blood sample competes with a known quantity of labelled antigen, in this case labelled AIDS env protein, for a limited quantity of antibody binding sites. Thus, the amount of labelled antigen bound to the antibody is inversely proportional to the amount of antigen in the sample. In another embodiment, an immunometric assay may be used wherein a labelled AIDS-env antibody is used. In such an assay, the amount of labelled antibody which complexes with the antigen-bound antibody is directly proportional to the amount of antigen (AIDS virus) in the blood sample. In a simple yes/no assay to determine whether the AIDS virus is present in blood, the solid support is tested to detect the presence of labelled antibody. In another embodiment, monoclonal antibodies to AIDS env protein may be used in an immunometric assay. Such monoclonal antibodies may be obtained by methods well known in the art, particularly the process of Milstein and Kohler reported in Nature 256, 495-497 (1975).

The immunometric assay method is as follows: Duplicate samples are run in which 100 µl of a suspension of antibody immobilized on agarose particles is mixed with 100 µl of serum and 100 µl of soluble ^{125}I-labelled antibody. This mixture is for specified times ranging from one quarter hour to twenty four hours. Following the incubation periods the agarose particles are washed by addition of buffer and then centrifuged. After removal of the washing liquid by aspiration, the resulting pellet of agarose particles is then counted for bound ^{125}I-labelled antibody. The counts obtained for each of the complexes can then be compared to controls.

While the invention has been described in terms of certain preferred embodiments, modifications obvious to one with ordinary skill in the art may be made without departing from the scope of the invention. For example, it is understood that the env AIDS DNAs described herein represent only the precise structure of two naturally occurring gene segments. It is expected that slightly modified alleles will be found encoding for similarly functioning proteins, and such gene segments and proteins are considered to be equivalents for the purpose of this invention. It is also suspected that other variants in addition to those described herein will be found and that the envelope protein of said variants will differ slightly. These variant envelope proteins are likewise considered within the scope of the invention. DNA having equivalent codons is considered within the scope of the invention, as are synthetic gene segments that encode homologous proteins of the viral envelope.

Various features of the invention are set forth in the following claims.

form. It is also foreseeable that monoclonal antibody affinity chromatography columns utilizing env AIDS polyclonal or monoclonal antibodies, could be used as an alternative to HPLC.

Example 5

Polyacrylamide gel electrophoresis and Western blot analysis

Cells were lysed by resuspending the cell pellets (approximately 10^8 cells) in TG buffer (10 mM Tris, pH 7.4, 10% glycerol), mixed with an equal volume of 2 x sample buffer [Laemmli, U.K., "Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4", Nature 227, 680-685 (1970)] and incubated at 95°C for five (5) minutes. Cell debris were pelleted by centrifugation and the cleared lysates were subjected to SDS-PAGE analysis [Laemmli, U.K., supra]. For Western blot analysis, the proteins from the acrylamide gel were electroblotted onto a 0.1 µm nitrocellulose membrane (Schleicher and Schuell) for 16 hr at 50V, in 12.5 mM Tris, 96 mM glycine, 20% methanol, 0.01% SDS at pH 7.5. Processing of the blot was carried out using the methods described by Towbin, H. et al. ["Electrophoretic Transfer of Proteins From Polyacrylamide Gels to Nitrocellulose Sheets: Procedure and Some Applications", Proc. Natl. Acad. Sci. U.S.A., 76, 4350-4354, (1979)]. For treatment with the human sera, the blots were incubated with a 1000 fold dilution of the sera in antibody buffer (20 mM sodium phosphate buffer, pH 7.5, containing 0.5 M NaCl, 1% BSA and 0.05% Tween 20) for 2-6 hr. The blots were then washed twice with phosphate buffered saline containing 0.05% Tween 20 and then incubated with 125-I-labelled Staphylococous aureus protein A for an additional period of 1 hr. The blot was then washed twice in PBS-Tween 20 buffer, dried and autoradiographed.

Example 6

Immunization with Env Protein of AIDS Virus

It is clear that in spite of the divergence observed between HTLVIII and ARV-2 sequences, the recombinant proteins synthesized by the bacteria react well with AIDS patients' sera derived from both geographical locations of the United States. One hundred percent (100%) of the AIDS patients' sera tested showed high reactivity (50 individual samples, 25 from the east coast of the United States and 25 from the west coast of the United States). Thus all the env proteins contain at least one conserved epitope. All of the human sera from AIDS patients tested contained antibodies to the env proteins of the instant invention. This strongly suggests that these env proteins with the conserved epitopes would be immunogenic in man.

It will be readily appreciated that the env proteins of the instant invention can be incorporated into vaccines capable of inducing protective immunity against the AIDS virus. By methods known in the art, the specific amino acids conprising the epitopes of the env protein may be determined. Peptides may then be synthesized, comprising an amino acid sequence corresponding to an epitope of an env AIDS protein either in monomeric or multimeric form. These synthetic peptides may then be incorporated into vaccines capable of inducing protective immunity against AIDS virus. Techniques for enhancing the antigenicity of such peptides include incorporation into a multimeric structure, binding to a highly immunogenic protein carrier, for example, keyhold limpet hemocyanin, or diphtheria toxoid, and administration in combination with adjuvants or any other enhancers of immune response. In addition, the vaccine composition may comprise antigens to provide immunity against other diseases in addition to AIDS.

An amino acid sequence corresponding to an epitope of an env protein either in monomeric or multimeric form (peptide) may be obtained by chemical synthetic means or by purification from biological sources including genetically modified microorganisms or their culture media. The peptide may be combined in an amino acid sequence with other peptides including fragments of other proteins, as for example, when synthesized as a fusion protein, or linked to other antigenic or non-antigenic peptides of synthetic or biological origin. The term "corresponding to an epitope of a env protein" will be understood to include the practical possibility that, in some instances, amino acid sequence variations of a naturally occurring peptide may be antigenic and confer protective immunity against AIDS infection. Possible sequence variations include, without limitation, amino acid substitutions, extensions, deletions, interpolations and combinations thereof. Such variations fall within the contemplated scope of the invention provided the peptide containing them is antigenic and antibodies elicited by such peptide cross-react with naturally occurring env protein or non-variant repeated peptides of env protein, to an extent sufficient to provide protective immunity when administered as a vaccine. Such vaccine compositions will be combined with a physiologically acceptable medium.

METArg
ValLysGluLysTyrGlnHisLeuTrpArgTrpGlyTrpArgTrpGlyThrMETLeuLeuGlyMETLeu
METIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHisSerLeuIle
GluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu
TrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyrIleLysLeuPheIleMETIleValGlyGlyLeu
ValGlyLeuArgIleValPheAlaValLeuSerValValAsnArgValArgGlnGlyTyrSerProLeu
SerPheGlnThrHisLeuProIleProArgGlyProAspArgProGluGlyIleGluGluGluGlyGly
GluArgAspArgAspArgSerIleArgLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArg
SerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuLeuLeuIleValThrArgIleValGluLeu
LeuGlyArgArgGlyTrpGluAlaLeuLysTyrTrpTrpAsnLeuLeuGlnTyrTrpSerGlnGluLeu
LysAsnSerAlaValSerLeuLeuAsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIle
GluValValGlnGluAlaTyrArgAlaIleArgHisIlePro ArgArgIleArgGlnGlyLeuGluArg
IleLeuLeu

3. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

Claims

1. An envelope protein of an acquired immune deficiency syndrome (AIDS) virus comprising the amino acid sequence:

METArg
ValLysGluLysTyrGlnHisLeuTrpArgTrpGlyTrpArgTrpGlyThrMETLeuLeuGlyMETLeu
METIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHisSerLeuIle
GluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu
TrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyrIleLysLeuPheIleMETIleValGlyGlyLeu
ValGlyLeuArgIleValPheAlaValLeuSerValValAsnArgValArgGlnGlyTyrSerProLeu
SerPheGlnThrHisLeuProIleProArgGlyProAspArgProGluGlyIleGluGluGluGlyGly
GluArgAspArgAspArgSerIleArgLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArg
SerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuLeuLeuIleValThrArgIleValGluLeu
LeuGlyArgArgGlyTrpGluAlaLeuLysTyrTrpTrpAsnLeuLeuGlnTyrTrpSerGlnGluLeu
LysAsnSerAlaValSerLeuLeuAsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIle
GluValValGlnGluAlaTyrArgAlaIleArgHisIleProArgArgIleArgGlnGlyLeuGluArg
IleLeuLeu

or fragments thereof.

2. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

5. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

METArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

6. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

25

METTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

7. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

40

METArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

8. An envelope protein as claimed in any one of claims 1 to 7 that is a homogeneous protein free of other AIDS viral proteins.

55

9. An expression vector comprising a gene coding for an envelope protein of an AIDS virus downstream of a promoter sequence enabling transcription, translation and thus expression of said envelope protein in a host cell.

ValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

4. An envelope protein of an AIDS virus according to claim 1 with the amino acid sequence:

CysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

TGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof.

12. An expression vector according to claim 9 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

ATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof.

13. An expression vector according to claim 9 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

10. An expression vector according to claim 9 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

GTGTGGAAGGAAGCA
ACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACA
CATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAAC
ATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAG
CCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACC
AATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACA
AGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAAT
GATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof.

11. An expression vector according to claim 9 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

ence of antibodies to the viral etiologic agent of AIDS which comprises mixing a composition containing an envelope protein of an AIDS virus as claimed in any one of claims 1 to 8 with a sample of human blood and determining whether said envelope AIDS protein binds to AIDS antibodies present in the blood sample.

30. A method according to claim 29 which comprises the use of the Western Blotting Analysis.

31. A method according to claim 29 which comprises the use of an ELISA-technique, wherein an envelope protein of an AIDS virus as claimed in any one of claims 1 to 8 is coated on a solid phase and contacted with the sample and after washing contacted with an enzyme-labeled non-human IgG.

32. A method according to claim 29, wherein the Double-Antigen-Method is used.

33. A method for the determination of AIDS virus, wherein antibodies against an envelope protein of an AIDS virus according to any one of claims 1 to 8 are used.

34. A method according to claim 33, wherein the antigen in the sample and a protein as claimed in any one of claims 1 to 8 in labeled form compete with an antibody against a protein as claimed in any one of claims 1 to 8.

35. A method according to claim 33, wherein a sandwich method is performed using two antibodies against a protein as claimed in any one of claims 1 to 8.

36. A method according to claim 35, wherein one antibody is on a solid phase and the other antibody is labeled.

37. A method according to claim 35, wherein two different monoclonal antibodies are used.

38. A vaccine eliciting immunity to AIDS comprising as an active ingredient a protein as claimed in any one of claims 1 to 8.

39. Antibodies raised against a protein as claimed in any one of claims 1 to 8.

40. The antibodies of claim 39 which are monoclonal antibodies.

41. The use of a protein as claimed in anyone of claims 1-8 for the preparation of a protective immunisation vaccine.

42. The use of a protein as claimed in any one of claims 1 to 8 for testing human blood for the presence of AIDS virus.

Claims for Austria:

1. A process for the preparation of an envelope protein of an acquired immune deficiency syndrome (AIDS) virus comprising:

transforming a host cell with an expression vector comprising a gene coding for an envelope protein of an AIDS virus downstream of a promoter sequence enabling transcription, translation and expression of said envelope protein in said host cell;

culturing said host cell so that said envelope protein of an AIDS virus is expressed; and

extracting and isolating said envelope protein of an AIDS virus.

2. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

ATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof.

14. An expression vector according to claim 9 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

ATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCT

15. An expression vector according to any one of claims 9 to 14 which is a plasmid capable of replication in gram-negative and/or gram-positive bacteria.

16. An expression vector according to claim 15 which is capable of replication in an E. coli strain.

17. An expression vector according to claim 15 which is capable of replication in a B. subtilis strain.

18. An expression vector according to claim 15 or 16 which is a member of the pEV/env family.

19. An expression vector according to claim 18 which is pEV1, -2, or -3/env 44-640.

20. An expression vector according to claim 18 which is pEV1, -2, or -3/env 205-640.

21. A transformant carrying an expression vector as claimed in any one of claims 9 to 20.

22. A transformant according to claim 21 which is an E. coli strain.

23. A transformant according to claim 22 which is an E. coli MC 1061 strain.

24. A transformant according to claim 21 which is a B. subtilis strain.

25. A transformant according to claim 21 which is a eucaryotic cell.

26. A method of producing an envelope protein of an acquired immune deficiency syndrome virus as claimed in any one of claims 1 to 8 comprising:

transforming a host cell with an expression vector as claimed in any one of claims 9 to 20;

culturing said host cell so that said AIDS env protein is expressed; and,

extracting and isolating said AIDS env protein.

27. A method according to claim 26 wherein the expression vector is pEV1, -2 or -3/env 44-640.

28. A method according to claim 26 wherein the expression vector is pEV1, -2 or -3/env 205-640.

29. A method of testing human blood for the pres-

ValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

is used.

4. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

CysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

METArg
ValLysGluLysTyrGlnHisLeuTrpArgTrpGlyTrpArgTrpGlyThrMETLeuLeuGlyMETLeu
METIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHisSerLeuIle
GluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu
TrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyrIleLysLeuPheIleMETIleValGlyGlyLeu
ValGlyLeuArgIleValPheAlaValLeuSerValValAsnArgValArgGlnGlyTyrSerProLeu
SerPheGlnThrHisLeuProIleProArgGlyProAspArgProGluGlyIleGluGluGluGlyGly
GluArgAspArgAspArgSerIleArgLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArg
SerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuLeuLeuIleValThrArgIleValGluLeu
LeuGlyArgArgGlyTrpGluAlaLeuLysTyrTrpTrpAsnLeuLeuGlnTyrTrpSerGlnGluLeu
LysAsnSerAlaValSerLeuLeuAsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIle
GluValValGlnGluAlaTyrArgAlaIleArgHisIleProArgArgIleArgGlnGlyLeuGluArg
IleLeuLeu

is used.

3. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

is used.

8. A process according to any one of claims 1 to 7 wherein the host cell is a bacterium.

9. A process according to claim 8 wherein the bacterium is E. coli.

10. A process according to claim 9 wherein the expression vector is a plasmid of the pEV/env family.

11. A process according to claim 10 wherein the plasmid is pEV1, -2, or -3/env 44-640.

12. A process according to claim 10 wherein the plasmid is pEV1, -2, or -3/env 205-640.

13. A process for the preparation of an expression vector comprising a gene coding for an envelope protein of an AIDS virus, which process comprises constructing an expression vector having an insertion site wherein said gene may be inserted which insertion site is downstream of a promoter sequence enabling transcription, translation and thus expression of said envelope protein in a host cell.

14. A process according to claim 13 characterized in that as a gene coding for an envelope protein of an AIDS virus a gene comprising the nucleotide sequence

GTGTGGAAGGAAGCA
ACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACA
CATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAAC
ATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAG
CCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACC
AATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACA
AGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAAT
GATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof is used.

15. A process according to claim 13 characterized in that as a gene coding for an envelope protein of an AIDS virus a gene comprising the nucleotide sequence

is used.

5. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

METArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

is used.

6. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

METTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

is used.

7. A process according to claim 1, characterized in that as a gene coding for an envelope protein of an AIDS virus a gene capable of effecting expression of the protein comprising the amino acid sequence of:

METArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSer

ATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof is used.

18. A process according to claim 13 characterized in that as a gene coding for an envelope protein of an AIDS virus a gene comprising the nucleotide sequence

ATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCT

or an equivalent thereof is used.

19. A process according to any one of claims 13 to 18 wherein the expression vector is a plasmid capable of replication in gram-negative bacteria.

20. A process according to claim 19 wherein the plasmid is capable of replication in an E. coli strain.

21. A process for the preparation of a transformant carrying an expression vector comprising a gene coding for an envelope protein of an AIDS virus, which process comprises transforming a microorganism with an expression vector obtained according to any one of claims 13 to 20 and cultivating the transformed microorganism.

22. A process according to claim 21 wherein the microorganism is an E. coli strain.

23. A process according to claim 22 wherein the microorganism is an E. coli MC 1061 strain.

24. A process of testing human blood for the presence of antibodies to the viral etiologic agent of AIDS which process comprises mixing a composition containing an evelope protein of an AIDS virus obtained according to any one of claims 1 to 12 with a sample of human blood and determining whether said envelope AIDS protein binds to AIDS antibodies present in the blood sample.

25. A process according to claim 24 which comprises the use of the Western Blotting Analysis.

26. A process according to claim 24 which comprises the use of an Elisa-technique, wherein an envelope protein of an AIDS virus obtained according to any one of claims 1 to 12 is coated on a solid phase and contacted with the sample and after washing contacted with an enzyme-labeled non-human IgG.

27. A process according to claim 24, wherein the Double-Antigen-Method is used.

28. A process for the determination of AIDS virus, wherein antibodies against an envelope protein of an AIDS virus obtained according to any one of claims 1 to 12 are used.

29. A process according to claim 28, wherein the

TGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof is used.

16. A process according to claim 13 characterized in that as a gene coding for an envelope protein of an AIDS virus a gene comprising the nucleotide sequence

ATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof is used.

17. A process according to claim 13 characterized in that as a gene coding for an envelope protein of an AIDS virus a gene comprising the nucleotide sequence

or an equivalent thereof.

39. An expression vector according to claim 37 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

```
                                                                TGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC
```

or an equivalent thereof.

40. An expression vector according to claim 37 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

```
   ATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC
```

antigen in the sample and a protein obtained according to any one of claims 1 to 12 in labeled form compete with an antibody against a protein obtained according to any one of claims 1 to 12.

30. A process according to claim 28, wherein a sandwich method is performed using two antibodies against a protein obtained according to any one of claims 1 to 12.

31. A method according to claim 30, wherein one antibody is on a solid phase and the other antibody is labeled.

32. A method according to claim 30, wherein two different monoclonal antibodies are used.

33. An envelope protein of an AIDS virus whenever prepared by a process as claimed in any one of claims 1 to 12.

34. An envelope protein of an AIDS virus according to claim 33 that is a homogeneous protein free of other AIDS viral proteins.

35. An expression vector comprising a gene coding for an envelope protein of an AIDS virus whenever prepared by a process as claimed in any one of claims 13 to 20.

36. A transformant carrying an expression vector comprising a gene coding for an envelope protein of an AIDS virus whenever prepared by a process as claimed in any one of claims 21 to 23.

37. An expression vector comprising a gene coding for an envelope protein of an AIDS virus downstream of a promoter sequence enabling transcription, translation and thus expression of said envelope protein in a host cell.

38. An expression vector according to claim 37 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

GTGTGGAAGGAAGCA
ACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACA
CATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAAC
ATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAG
CCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACC
AATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACA
AGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAAT
GATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCC
TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG
TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA
TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG
GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC
AATACAAGAAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA
AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC
AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA
ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT
ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC
CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

54. The use of a protein as claimed in anyone of claims 1-8 for the preparation of a protective immunisation vaccine.

or an equivalent thereof.

41. An expression vector according to claim 37 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

ATGTATGCCCCTCCCATC
AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC
AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGC

or an equivalent thereof.

42. An expression vector according to claim 37 wherein said gene coding for an envelope protein of an AIDS virus is a gene comprising the nucleotide sequence:

ATGAGGGACAATTGGAGAAGTGAATTATATAAA
TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA
GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC
GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT
TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG
GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA
AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG
AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCT

43. An expression vector according to any one of claims 37 to 42 which is a plasmid capable of replication in gram-negative bacteria.

44. An expression vector according to claim 43 which is capable of replication in an E. coli strain.

45. An expression vector according to claim 43 or 44 which is a member of the pEV/env family.

46. An expression vector according to claim 45 which is pEV1, -2, or -3/env 44-640.

47. An expression vector according to claim 45 which is pEV1, -2, or -3/env 205-640.

48. A transformant carrying an expression vector as claimed in any one of claims 37-47.

49. A transformant according to claim 48 which is an E. coli strain.

50. A transformant according to claim 49 which is an E. coli MC 1061 strain.

51. Antibodies raised against a protein obtained according to any one of claims 1 to 12 and 33 and 34.

52. The antibodies of claim 51 which are monoclonal antibodies.

53. A vaccine eliciting immunity to AIDS comprising as an active ingredient a protein obtained according to any one of claims 1 to 12 and 33 and 34.

FIGURE 1

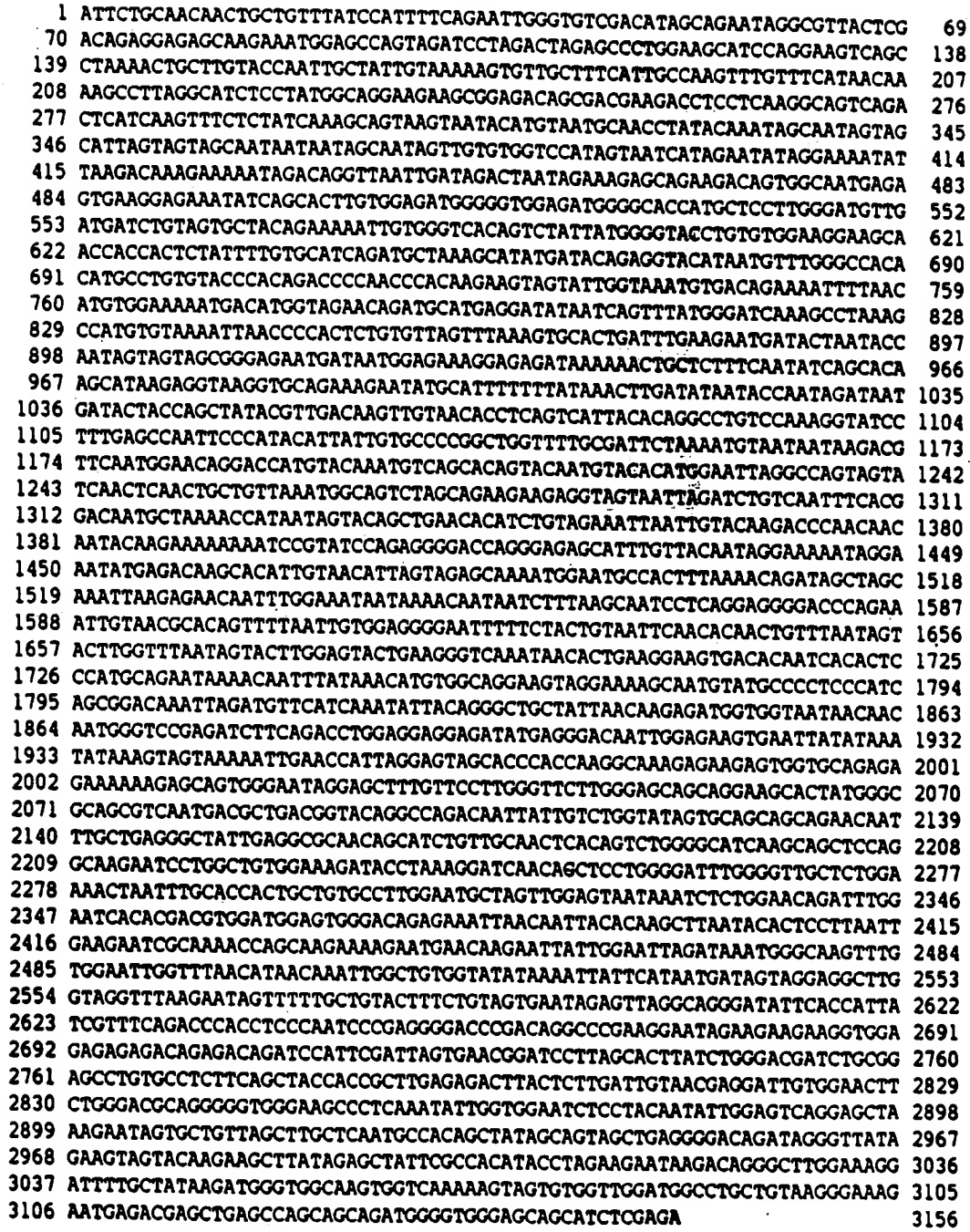

1 ATTCTGCAACAACTGCTGTTTATCCATTTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCG 69
70 ACAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGC 138
139 CTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAA 207
208 AAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGACCTCCTCAAGGCAGTCAGA 276
277 CTCATCAAGTTTCTCTATCAAAGCAGTAAGTAATACATGTAATGCAACCTATACAAATAGCAATAGTAG 345
346 CATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATAT 414
415 TAAGACAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAGAGCAGAAGACAGTGGCAATGAGA 483
484 GTGAAGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAGATGGGGCACCATGCTCCTTGGGATGTTG 552
553 ATGATCTGTAGTGCTACAGAAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCA 621
622 ACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACA 690
691 CATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAAC 759
760 ATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAG 828
829 CCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACC 897
898 AATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACA 966
967 AGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAAT 1035
1036 GATACTACCAGCTATACGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCC 1104
1105 TTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACG 1173
1174 TTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTA 1242
1243 TCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGTCAATTTCACG 1311
1312 GACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAAC 1380
1381 AATACAAGAAAAAAAATCCGTATCCAGAGGGGACCAGGGAGAGCATTTGTTACAATAGGAAAAATAGGA 1449
1450 AATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGCCACTTTAAAACAGATAGCTAGC 1518
1519 AAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAATCCTCAGGAGGGGACCCAGAA 1587
1588 ATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT 1656
1657 ACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACTC 1725
1726 CCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC 1794
1795 AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC 1863
1864 AATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAA 1932
1933 TATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGA 2001
2002 GAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGC 2070
2071 GCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT 2139
2140 TTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAG 2208
2209 GCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGA 2277
2278 AAACTAATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTTGG 2346
2347 AATCACACGACGTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATT 2415
2416 GAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG 2484
2485 TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTG 2553
2554 GTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTGTAGTGAATAGAGTTAGGCAGGGATATTCACCATTA 2622
2623 TCGTTTCAGACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGA 2691
2692 GAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTGCGG 2760
2761 AGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTT 2829
2830 CTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAATCTCCTACAATATTGGAGTCAGGAGCTA 2898
2899 AAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCTATAGCAGTAGCTGAGGGGACAGATAGGGTTATA 2967
2968 GAAGTAGTACAAGAAGCTTATAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGG 3036
3037 ATTTTGCTATAAGATGGGTGGCAAGTGGTCAAAAAGTAGTGTGGTTGGATGGCCTGCTGTAAGGGAAAG 3105
3106 AATGAGACGAGCTGAGCCAGCAGCAGATGGGGTGGGAGCAGCATCTCGAGA 3156

FIGURE 2 (3 pages)

```
          1                                                                  50
HXB-3     MRVKEK-----YQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATT
BH-10
BH-8                                                          F
LAV                              K        I
ARV-2      K  --GTRRN            ----- -                     -          --

         51                                                             100
HXB-3     TLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDM
BH-10
BH-8
LAV
ARV-2               R                          G           N

        101                                                             150
HXB-3     VEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSS-----SGRMIME
BH-10
BH-8
LAV                                          G A       NTNSS  E M
ARV-2          Q                        T N  G A       NWKEEI------

        151                                                             200
HXB-3     KGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDND--TTSYTLTS---CNTSV
BH-10
BH-8                    K
LAV
ARV-2                T    D I   N L RN  VV    AST  N  NYRLIH  R

        201                                                             250
HXB-3     ITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHG
BH-10
BH-8
LAV                                      A
ARV-2                       T              K

        251                                                             300
HXB-3     IRPVVSTQLLLNGSLAEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPN
BH-10                              A              Q
BH-8                                             D
LAV                                A              Q
ARV-2          I                   D   N          E  A
```

```
      301                                                        350

HXB-3   NNTRKKIRIQRGPGRAFVTIGKIGNMRQ-AHCNISRAKWNATLKQIASKLR
BH-10        S                                        N     D
BH-8                                                        D
LAV          S
ARV-2        S Y --      H T R IGDIRK        Q  N  E  VK


      351                                                        400

HXB-3   EQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTW
BH-10
BH-8
LAV
ARV-2           V N          M     R       T     N  -RLNH


      401                                                        450

HXB-3   STEGSNNTEGSDTITLPCRIKQFINMWQEVGKAMYAPPISGQIRCSSNIT
BH-10     K                  I
BH-8      K                  I
LAV
ARV-2   -   --- K N   I       I              G   S


      451                                                        500

HXB-3   GLLLTRDGG-NNNNGSEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTK
BH-10           - S  E
BH-8            - S  E
LAV             -                        --
ARV-2           T VT DT V                   I      I


      501                                                        550

HXB-3   AKRRVVQREKRAVGI-GALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQ
BH-10                 -
BH-8                  -
LAV                   -                R
ARV-2                 V  M             V L


      551                                                        600

HXB-3   QQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSG
BH-10            G
BH-8
LAV
ARV-2                                V       R
```

```
      601                                                        650

HXB-3   KLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLIEESQ
BH-10                            NM
BH-8                             NM
LAV                              NM
ARV-2                      D  DNM   Q E   D   NT YT

      651                                                        700

HXB-3   NQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFA
BH-10
BH-8
LAV                                      I
ARV-2                          S         I

      701                                                        750

HXB-3   VLSVVNRVRQGYSPLSFQTHLPIPRGPDRPEGIEEEGGERDRDRSIRLVN
BH-10
BH-8       I                N
LAV        I                T
ARV-2      I              R V       D              V   D

      751                                                        800

HXB-3   GSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNLL
BH-10
BH-8
LAV
ARV-2    F     E          R        AA T  I  H          S

      801                                                        850

HXB-3   QYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQEAYRAIRHIPRRIRQG
BH-10                                  G
BH-8              N                 L  A
LAV                                    G C
ARV-2      I         W        T        A R     L  H

      851  856

HXB-3   LERILL
BH-10
BH-8
LAV
ARV-2      L
```

" - " designates a deletion of one amino acid. An empty space denotes identity with HXB-3 sequence.

Figure 3

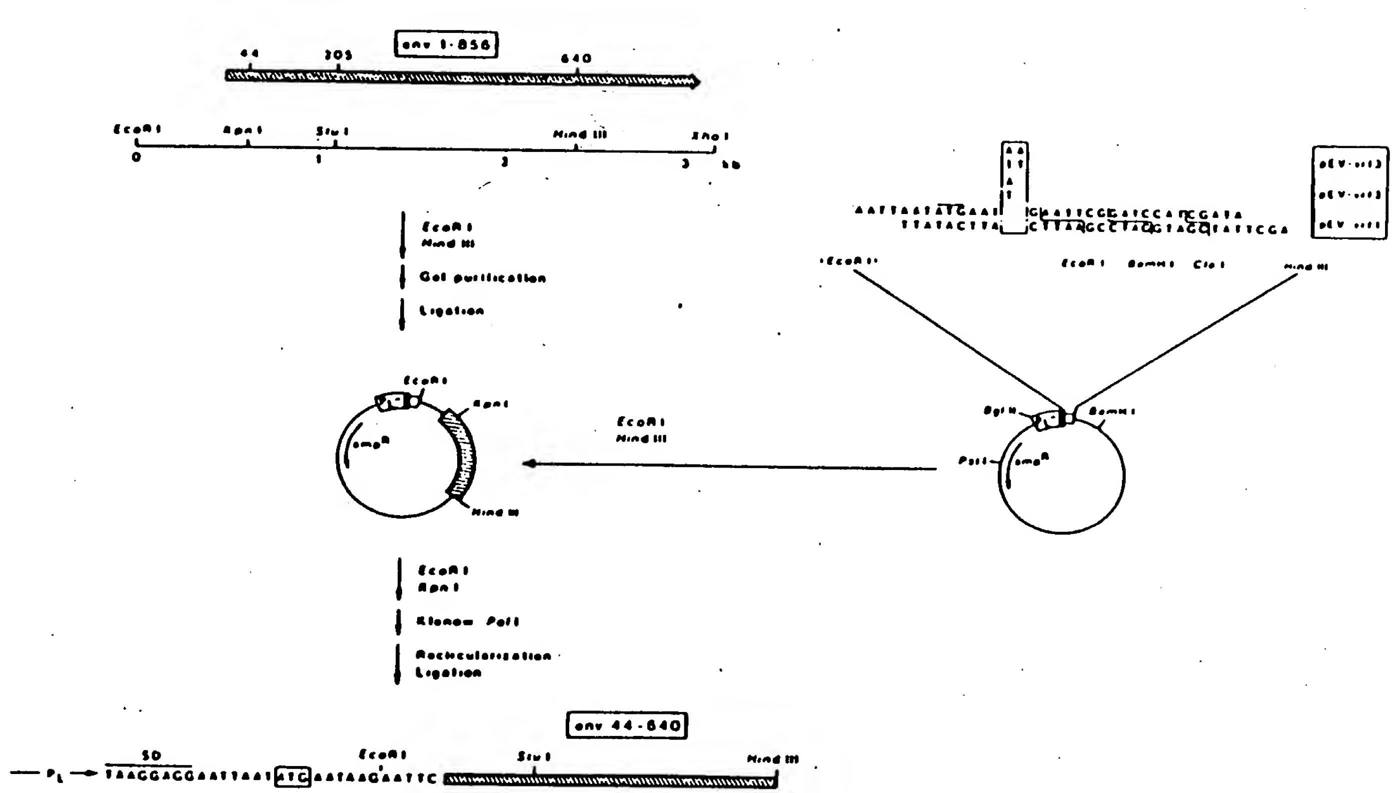

env 1-856
44
205
640
EcoRI
KpnI
StuI
HindIII
XhoI
0
1
2
3
kb
EcoRI
HindIII
Gel purification
Ligation
ampR
EcoRI
HindIII
EcoRI
KpnI
Klenow PolI
Recircularization
Ligation
env 44-640
SD
TAAGGAGGAATTAATATGAATAAGAATTC
PL
BglII
BamHI
PstI
EcoRI
BamHI
ClaI
HindIII

Figure 4

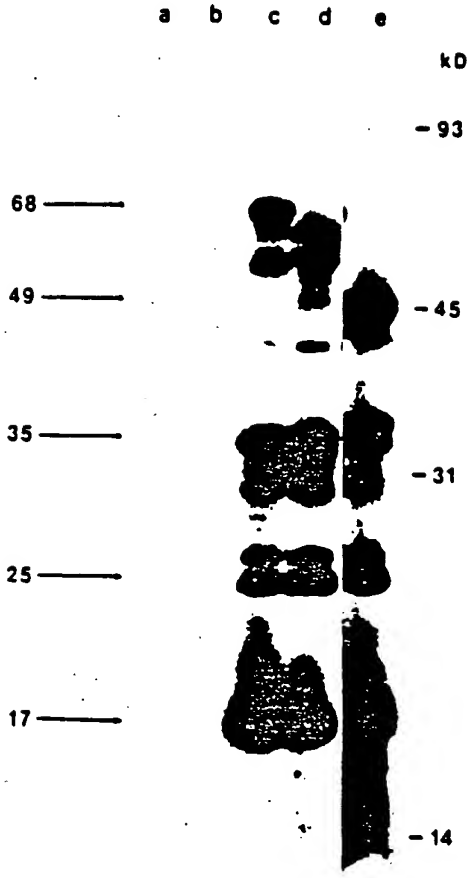
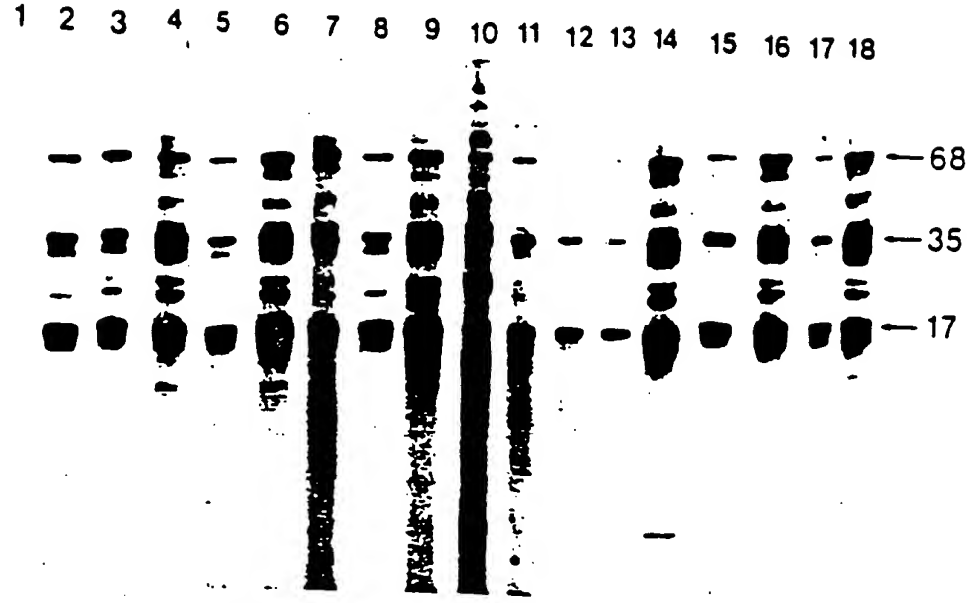


Figure 5


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
68
35
17

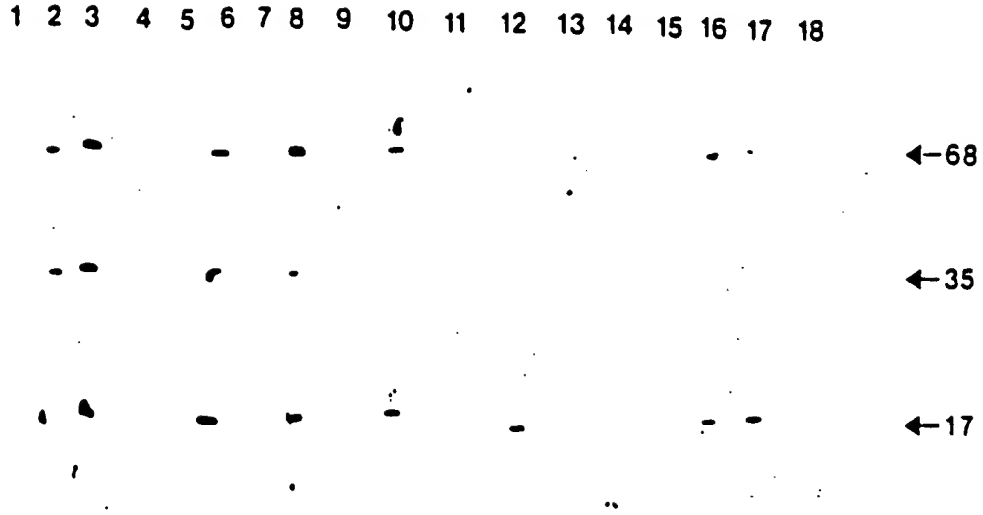

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
68
35
17

FIGURE 6A

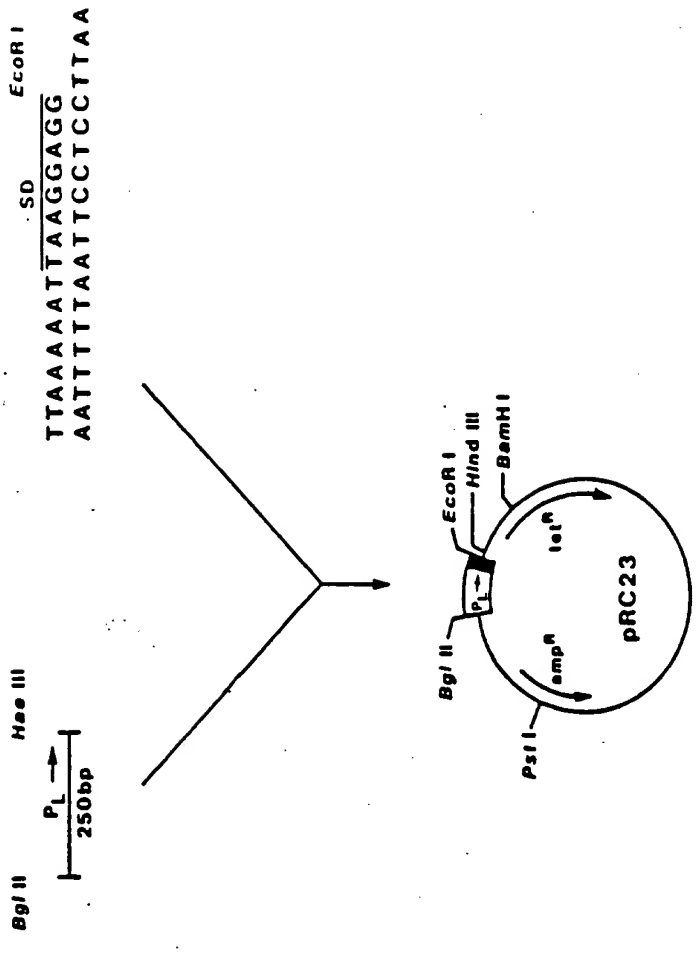
METArg
ValLysGluLysTyrGlnHisLeuTrpArgTrpGlyTrpArgTrpGlyThrMETLeuLeuGlyMETLeu
METIleCysSerAlaThrGluLysLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAla
ThrThrThrLeuPheCysAlaSerAspAlaLysAlaTyrAspThrGluValHisAsnValTrpAlaThr
HisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsn
METTrpLysAsnAspMETValGluGlnMETHisGluAspIleIleSerLeuTrpAspGlnSerLeuLys
ProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr
AsnSerSerSerGlyArgMETIleMETGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThr
SerIleArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsn
AspThrThrSerTyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSer
PheGluProIleProIleHisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThr
PheAsnGlyThrGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValVal
SerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValValIleArgSerValAsnPheThr
AspAsnAlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsnAsn
AsnThrArgLysLysIleArgIleGlnArgGlyProGlyArgAlaPheValThrIleGlyLysIleGly
AsnMETArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAlaThrLeuLysGlnIleAlaSer
LysLeuArgGluGlnPheGlyAsnAsnLysThrIleIlePheLysGlnSerSerGlyGlyAspProGlu
IleValThrHisSerPheAsnCysGlyGlyGluPhePheTyrCysAsnSerThrGlnLeuPheAsnSer
ThrTrpPheAsnSerThrTrpSerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeu
ProCysArgIleLysGlnPheIleAsnMETTrpGlnGluValGlyLysAlaMETTyrAlaProProIle
SerGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsnAsnAsn
AsnGlySerGluIlePheArgProGlyGlyGlyAspMETArgAspAsnTrpArgSerGluLeuTyrLys
TyrLysValValLysIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValValGlnArg
GluLysArgAlaValGlyIleGlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMETGly
AlaAlaSerMETThrLeuThrValGlnAlaArgGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsn
LeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGln
AlaArgIleLeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
LysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGlnIleTrp
AsnHisThrThrTrpMETGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHisSerLeuIle
GluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu
TrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyrIleLysLeuPheIleMETIleValGlyGlyLeu
ValGlyLeuArgIleValPheAlaValLeuSerValValAsnArgValArgGlnGlyTyrSerProLeu
SerPheGlnThrHisLeuProIleProArgGlyProAspArgProGluGlyIleGluGluGluGlyGly
GluArgAspArgAspArgSerIleArgLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArg
SerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuLeuLeuIleValThrArgIleValGluLeu
LeuGlyArgArgGlyTrpGluAlaLeuLysTyrTrpTrpAsnLeuLeuGlnTyrTrpSerGlnGluLeu
LysAsnSerAlaValSerLeuLeuAsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIle
GluValValGlnGluAlaTyrArgAlaIleArgHisIleProArgArgIleArgGlnGlyLeuGluArg
IleLeuLeu

FIGURE 6B

AMINO ACID DISTRIBUTION
OF AIDS ENV PROTEIN

Name	Number of Residues
A Alanine	47
B Aspartic Acid-Asparagine	0
C Cysteine	21
D Aspartic Acid	27
E Glutamic Acid	49
F Phenylalanine	26
G Glycine	58
H Histidine	14
I Isoleucine	63
K Lysine	44
L Leucine	83
M Methionine	17
N Asparagine	60
P Proline	29
Q Glutamine	42
R Arginine	52
S Serine	57
T Threonine	60
V Valine	56
W Tryptophan	31
Y Tyrosine	20
Z Glutamine-Glutamic Acid	0

Figure 7


TTAAAAATTAAGGAGG
AATTTTTAATTCCTCCTTAA
SD
EcoR I
Hae III
Bgl II
PL
250bp
EcoR I
Hind III
BamH I
Bgl II
PL
lacI^q
ampR
pRC23
Pst I

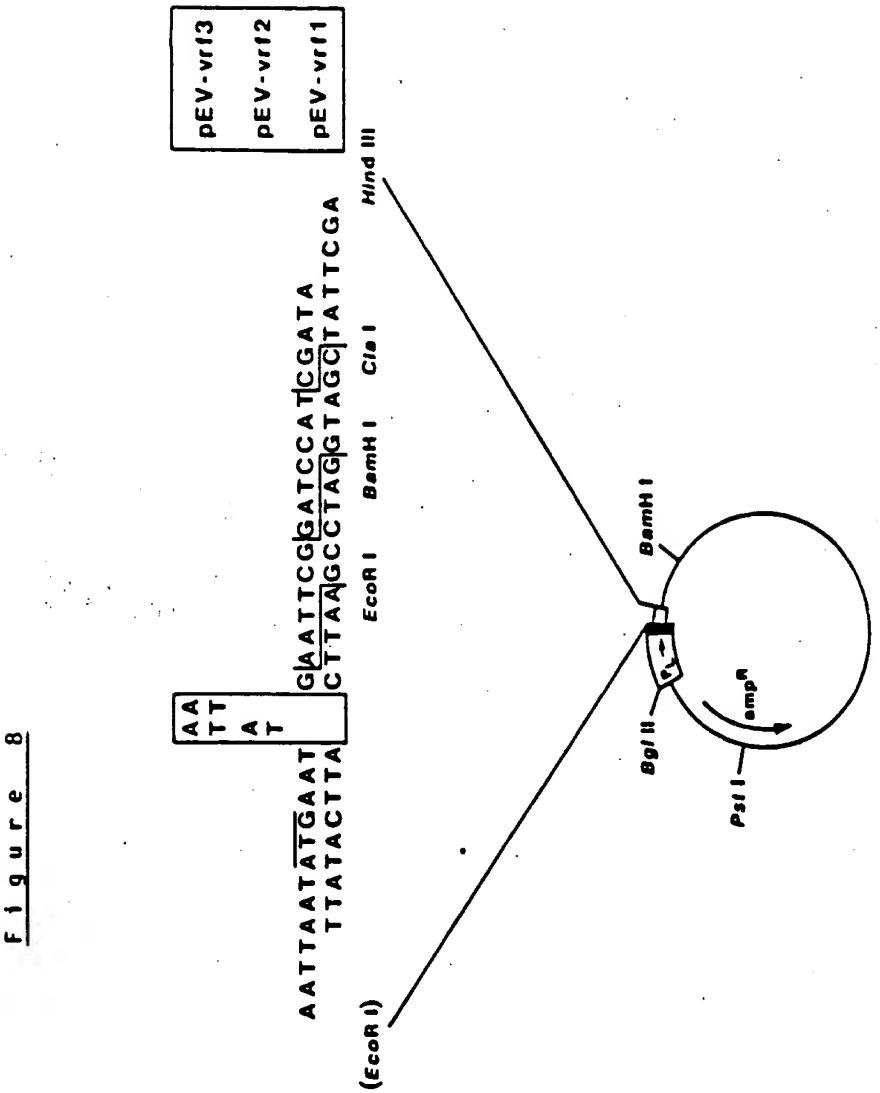

pEV-vrf3
pEV-vrf2
pEV-vrf1
AA
TT
A
T
AATTAATATGAAT
TTATACTTA
GAATTCGGATCCATCGATA
CTTAAGCCTAGGTAGCTATTCGA
EcoR I
BamH I
Cla I
Hind III
(EcoR I)
BamH I
P
Bgl II
ampR
Psi I

Figure 8

European Patent Office

EUROPEAN SEARCH REPORT

Application number

EP 86 10 5371

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl 4)
X,Y	SCIENCE, vol. 228, no. 4695, 5th April 1985, pages 93-96; N. CHANG et al.: "Expression in Escherichia coli of open reading frame gene segments of HTLV-III" * Whole article *	1-51	C 07 K 13/00 C 12 N 15/00 C 12 P 21/02 C 12 P 21/00 G 01 N 33/569 A 61 K 39/21
Y	NATURE, vol. 313, no. 6002, 7th February 1985, pages 450-458; M.A. MUESING et al.: "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus" * Whole article *	1,2	
A	IDEM	3-28	
Y	NATURE, vol. 313, no. 6000, 24th January 1985, pages 277-284, London, GB; L. RATNER et al.: "Complete nucleotide sequence of the AIDS virus, HTLV-III" * Whole article *	1,2	TECHNICAL FIELDS SEARCHED (Int. Cl.4) C 12 N G 01 N C 12 P A 61 K
A	IDEM	3-28	

-- -/-

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	28-06-1986	OSBORNE H.H.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO Form 1503 03 82

European Patent Office

EUROPEAN SEARCH REPORT

Application number

EP 86 10 5371

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 4)
Y	SCIENCE, vol. 226, no. 4679, 7th December 1984, pages 1165-1171; G.M. SHAW et al.: "Molecular characterization of human T-cell leukemia (lymphotropic) virus type III in the acquired immune deficiency syndrome" * Whole article *	1,2	
A	IDEM	3-28	
P,Y	EP-A-0 173 529 (THE UNITED STATES OF AMERICA) * Whole document *	1,2	
A		3-28, 33-51	
P,Y	EP-A-0 152 030 (JURIDICAL FOUNDATION JAPANESE FOUNDATION FOR CANCER RESEARCH) * Whole document *	1,2	TECHNICAL FIELDS SEARCHED (Int. Cl. 4)
A		3-51	
E	EP-A-0 181 150 (CHIRON CORP.) * Whole document *	1-52	
A	WO-A-8 404 327 (PRESIDENT & FELLOWS OF HARVARD COLLEGE) * Abstract; pages 1-14 *	24-34, 51,52	

-/-

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	28-06-1986	OSBORNE H.H.

EPO Form 1503 03 82

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document